

## KOREAN PATENT ABSTRACTS

(11)Publication number: **10-2003-0024309 A**

(43)Date of publication of application: **26.03.2003**

(51)Int. Cl.

**G06F 17 /00**

(21)Application number: **10-2001-0057389**

(22)Date of filing: **17.09.2001**

(71)Applicant: **JUNG, SANG HO**

(72)Inventor: **JUNG, SANG HO**

---

### (54) **PROGRAM FOR BLOCKING SLANG ON WEB BULLETIN BOARD**

(57) Abstract:

PURPOSE: A program for blocking slang on a web bulletin board is provided to transfer an upright meaning to other Netizen while protecting freedom for expressing a language on the Internet and to enable a web manager to efficiently manage the bulletin board, a monitoring function and an operation time setting function.

CONSTITUTION: A bulletin list after an input date of the latest processed bulletin is read by using a select query through a JDBC(Java DataBase Connectivity) interface of a bulletin board RDBMS(Relational DataBase Management System). A text of the bulletin is passed over to a tokenizer in order. The tokenizer divides the text into each token by scanning the text through a defined special character set. A slang extracting module finds a key of a slang list loaded on a main memory by using the first phoneme of each token. The data matched with the token in a pattern is searched from a specific list on the found slang list by using a binary search technique. If the pattern-matched token is found out, a substituted text is obtained from the tokenizer. The text of the bulleting including the found slang is updated by the substituted text.

(19) 대한민국특허청(KR)  
(12) 공개특허공보(A)

(51) Int. Cl. 7  
G06F 17/00

(11) 공개번호 특2003-0024309  
(43) 공개일자 2003년03월26일

(21) 출원번호 10-2001-0057389  
(22) 출원일자 2001년09월17일

(71) 출원인 정상호  
울산광역시 남구 무거1동 1240-7 대원빌라 301

(72) 발명자 정상호  
울산광역시 남구 무거동 1240-7 대원빌라 301

심사청구 : 있음

(54) 웹게시판 비어와 저속어 처리 프로그램

요약

본 아이디어는 갈수록 심해져 가는 인터넷의 비어, 저속어에 대해 기존의 방법과 다른 관리자중심의 처리 방안을 발명하였다.

기존 게시판의 비속어 처리프로그램은 비속어를 발견하면 입력을 할 수 없도록 차단하는 차단식 프로그램이었다.

이런 프로그램은 사이버 상의 의사표현의 자유를 차단한다. 또한 어떤 단어의 경우는 비속어가 아닌데도 차단되어 입력을 원천봉쇄하기도 한다.

그래서 비속어를 차단하지 않고 처리해 주며 신생 비속어도 처리를 쉽게 해 주며 검출에 유연성을 제공할 수 있는 프로그램을 개발하였다.

이프로그램을 개발 하므로 인해 인터넷에 동적인 콘텐츠를 제공하는 게시판의 비어, 저속어 등을 효율적으로 관리할 수 있다.

대표도

도 1

도 2

색인어

비어, 저속어, 웹크리너, 게시판관리, 네티켓, 언어폭력

명세서

도면의 간단한 설명

\* 프로그램명: 웹크리너

도1은 프로그램 전체 구성도

도2 는 비속어 처리모듈 프로그램

발명의 상세한 설명

발명의 목적

발명이 속하는 기술 및 그 분야의 종래기술

본발명은 인터넷의 비어와 저속어에 대한 관리의 어려움에 대해 효율적으로 관리 할 수 있는 비어, 저속어의 제거 프로그램에 관한것이다.

과거 인터넷 게시판의 비어 및 저속어의 처리 방식은 사용자가 입력 할 때 속어에 대한 차단으로 인하여 네티즌은 욕구 충족이 되지 않아 새로운 저속어(욕)을 만들게 하였다.

또한 인터넷이 발전함에 따라 일부 네티즌의 심한 저속어 표현에 대해 방어를 위한 프로그램으로 인해 네티즌들의 과감한 욕만들기가 성행하고 있는 실정이며, 게시판의 저속어를 차단하기 위하여 실제 언어로 사용되는 단어조차 차단하는 경우가 발생되어 네티즌의 언어표현의 자유를 막고 있다.

본 발명은 기존의 차단 방법이 아닌 네티즌의 언어표현의 자유를 지켜 주면서 게시판의 관리 모듈로 타 네티즌에게는 정화된 언어를 보여 주므로 인해 2중의 효과를 가져 올 수 있는 프로그램이다.

또한 웹마스터의 효율적인 관리를 위하여 여러 기능을 추가 하였다.

기존 프로그램과 차별된 것은 차단이 아닌 사용자가 저속어를 올린후 타 사용자가 그내용을 클릭 하였을때는 저속어는 치환 및 삭제 되면서 올바른 표현만 보여주는 방식이다.

발명이 이루고자 하는 기술적 과제

본 발명은 인터넷의 언어 표현의 자유를 지켜주면서 타 네티즌에게 올바른 뜻을 전달 될 수 있도록 제작된 프로그램이다.

또한 웹관리자가 게시판의 관리를 효율적으로 관리 할 수 있도록 치환기능, 삭제기능, 업로드 기능등을 추가 하였으며 모니터링 기능, 운영 시간 설정기능등을 원활히 관리 할수 있도록 하였다.

발명의 구성 및 작용

Webcleaner는 프로그래밍 모델 중MVC(Model-View- Controller) 디자인 패턴을 따라 개발하였으며 DB에 접근방법은 DB의 종류에 상관없는 JDBC(Java Database Connectivity)를 사용했다. Webcleaner는 Java언어로 작성되고 JDBC를 사용하므로 어떤 플랫폼에서든지 실행되며 또한 어떤 RDBMS(Relational DataBase Management Systems)에서든지 JDBC 드라이버가 있다면 연결되며 Java실행환경과 JSP/서블릿 엔진이 있으면 작동할 수 있다.

도1은 시스템의 구성도 이다. 그림의 Web server와 Web application server부를 보면 Http Request를 서버로부터 받아서 자바 실행환경의 비즈니스 로직부분을 호출하며 Controller의 역할을 하고 있고, 비즈니스 로직에서는 백엔드시스템과 통신한 결과를 Model인 자바빈즈에 넘겨준다. 그러면 JSP에서는 자바빈즈의 내용을 화면에 보여주는 View의 역할을 하여 Response를 보낸다. 백엔드시스템에는 게시판DB와 비속어리스트 파일, 결과를 저장하는 파일, 비속어를 처리할 게시판리스트 정보를 가지고 있는 게시판 정보파일이 있다. 특정 게시판에 대한 비속어처리를 시작하라는 Request가 있을 때 비즈니스 로직부의 cleaner 쓰레드 중 그 게시판에 관한 쓰레드를 시작시키게 된다. cleaner 쓰레드는 자바가상머신에서 멀티쓰레드로 실행되면서 주기억장치의 비속어리스트를 공유하므로 효율적이며 각 게시판에 대한 동시수행이 가능하다.

도 2는 게시판 RDBMS로 부터 본문 텍스트들을 얻어서 비속어를 추출하여 치환된 텍스트를 얻는 과정인데 순서대로 설명하면 다음과 같다.

- (1) 게시판RDBMS에 JDBC인터페이스를 사용하여 select 질의를 주어 가장마지막 처리한 게시물 의 입력날짜 이후의 게시물 리스트를 읽어온다.
- (2) 게시물의 본문텍스트를 차례로 토큰나이저에 넘겨준다. 그러면 토큰나이저가 정의된 특수문자세트에 의해 본문텍스트를 한번 스캐닝하여 단어(토큰)별로 나누게 하고 각 토큰의 본문텍스트 상의 인덱스를 기록해 둔다.
- (3) 비속어 추출모듈에서는 토큰나이저에게 토큰을 차례로 넘겨줄 것을 요청하여 각 토큰의 첫음절 음소에 의해 주기억장치에 올라와 있는 비속어 리스트의 키를 결정하여 찾아간다.
- (4) 찾아간 비속어 리스트 상의 특정리스트에서 토큰과 패턴이 매치되는 데이터가 있는지를 binary search기법으로 찾는다.
- (5) 이렇게 해서 본문텍스트의 끝까지 찾아서 비속어리스트와 패턴이 매치되는 토큰이 발견되었다면 토큰나이저에게 비속어 부분을 치환시킨 텍스트를 요청하여 치환텍스트를 얻어낸다. 토큰나이저는 토큰나이징 할 때 본문텍스트 상의 토큰의 인덱스를 기록해 두었으므로 치환텍스트를 만들어 낼 수 있다.
- (6) 게시판 RDBMS에 JDBC인터페이스를 사용하여 비속어가 발견된 게시물의 본문텍스트를 치환텍스트로 update하는 SQL문을 실행시켜 치환시킨다.
- (7) 치환시킨 날짜별로 치환시킨 게시물에 대한 정보를 결과파일에 저장한다.

본 발명의 전체 적인 구성은 저속어 분석 모듈(module), 설치모듈, 시스템 설정 모듈, 관리모듈로 나눌 수 있다.

#### 발명의 효과

이상과 같이 저속어 및 비어 처리 제거기는 웹관리자의 관리의 편리성과 네티즌의 심리적인 요인을 분석하여 차단이 아닌 제거로 인하여 인터넷 네티켓을 지켜나가는데 현실적인 프로그램이 될 수 있다.

본 발명의 결과를 정리하면 아래와 같다.

XX로 치환후	치환 전
X X~~ X X~!!	ㅁ 1~~ ㅂ ㅈ~!!
쌔쌔먹을 년이다	쌔쌔먹을 년이다
XXXXXXXXXXXX는	미원개첩창는
2001년	2001년
X	개
X	새
X	끼
뜬	뜬

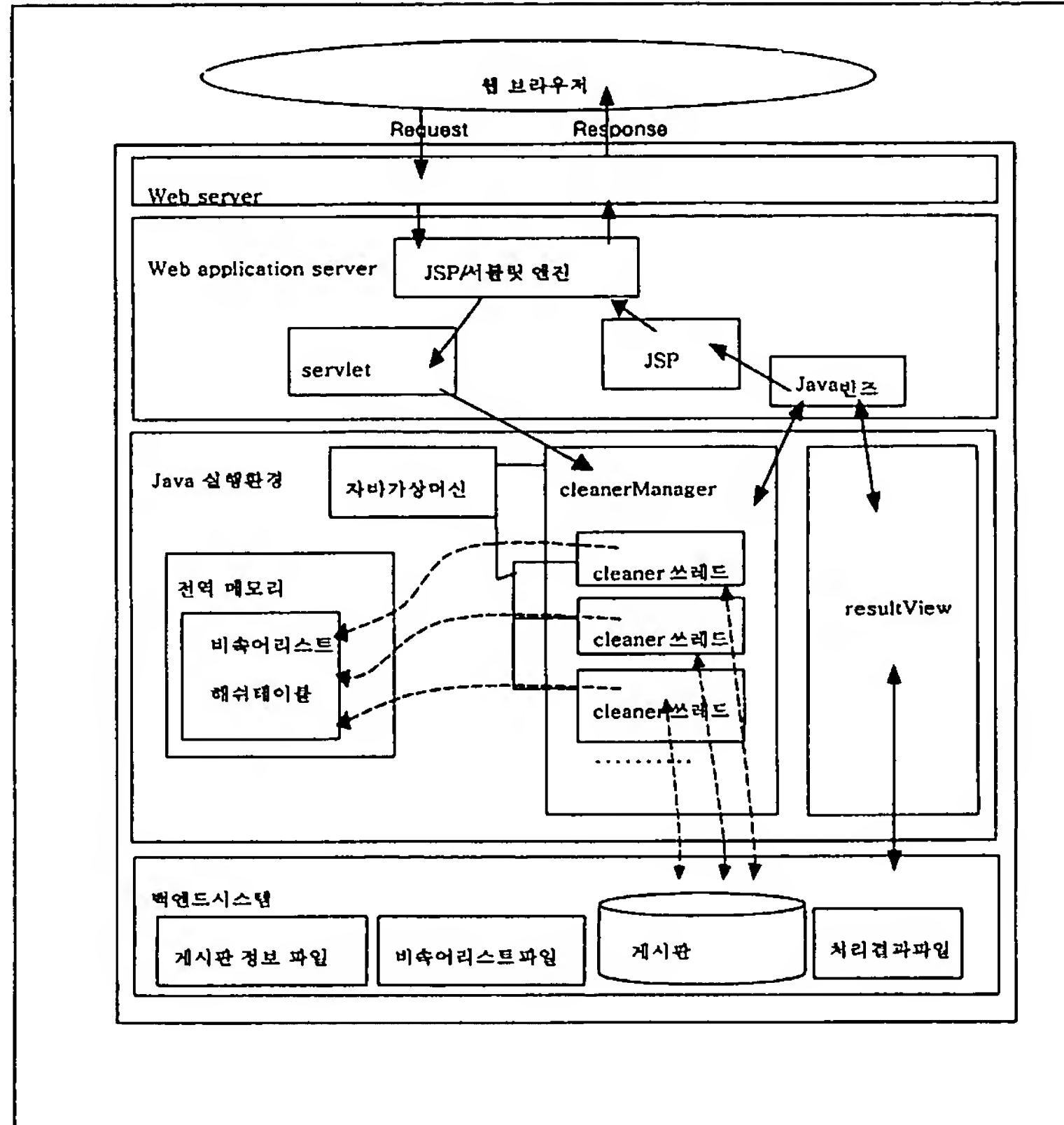
(57) 청구의 범위

청구항 1.

인터넷의 비어 및 저속어에 대한 효율적인 게시판 관리를 위하여 사용자가 속어등을 입력하는 경우 비어 및 저속어 분석기에 의하여 차단이 아닌 제거하는 프로그램

도면

도면 1



도면 2

